IN THE CLAIMS:

1.   (currently amended) A data table compressor, comprising:

a table modeller that discovers at least one model of data mining models with guaranteed error bounds of at least one attribute in a data table in terms of other attributes in different columns of said data table; ~~and~~

a model selector, associated with said table modeller, that selects a subset of said at least one model to form a basis upon which to compress said data table to form a compressed data table; and

a row aggregator that employs said selected subset from said model selector to improve a compression ratio of said compressed data table via row-wise clustering.

2.   (previously presented) The data table compressor as recited in Claim 1 wherein said table modeller employs classification and regression tree data mining models to model said at least one attribute.

3.   (previously presented) The data table compressor as recited in Claim 1 wherein said model selector employs a Bayesian network built on said at least one attribute to select relevant models for table compression.

4.   (previously presented) The data table compressor as recited in Claim 2 wherein construction of said models uses integrated building and pruning to exploit specified error bounds and decrease model construction time.

5.   (previously presented) The data table compressor as recited in Claim 1 wherein said table modeller employs a selected one of a constraint-based and a scoring-based method to generate said at least one model.

2

6.   (previously presented) The data table compressor as recited in Claim 1 wherein said model selector selects said subset based upon a compression ratio and an error bound specific for each attribute of said data table.

7.   (previously presented) The data table compressor as recited in Claim 2 wherein values for said at least one attribute are represented in said compressed data table by at least one of said classification and regression tree data mining models and are not explicitly stored therein.

8.   (previously presented) The data table compressor as recited in Claim 1 wherein said model selector selects said subset using a model built on attributes of said data table by a selected one of:

repeated calls to a maximum independent set solution algorithm, and

a greedy search algorithm.

9.   (currently amended) A method of compressing a data table, comprising:

discovering at least one model of data mining models with guaranteed error bounds of at least one attribute in said data table in terms of other attributes in different columns of said data table; and

selecting a subset of said at least one model to form a basis upon which to compress said data table; and

employing said selected subset to improve a compression ratio of said compressed data table via row-wise clustering.

10.   (original) The method as recited in Claim 9 wherein said discovering comprises employing classification and regression tree data mining models to model said at least one attribute.

11.   (original) The method as recited in Claim 9 wherein said discovering comprises employing a Bayesian network built on said at least one attribute to select relevant models for table compression.

12.   (original) The method as recited in Claim 10 further comprising using integrated building and pruning to exploit specified error bounds and decrease model construction time.

13.   (original) The method as recited in Claim 9 wherein said discovering comprises employing a selected one of a constraint-based and a scoring-based method to generate said at least one model.

14.   (original) The method as recited in Claim 9 wherein said selecting comprises selecting said subset based upon a compression ratio and an error bound specific for each attribute of said data table.

15.   (original) The method as recited in Claim 9 wherein said selecting is NP-hard.

16.   (original) The method as recited in Claim 9 wherein said selecting comprises selecting said subset using a model built on attributes of said data table by a selected one of:

repeated calls to a maximum independent set solution algorithm, and

a greedy search algorithm.

17.   (currently amended) A database management system, comprising:

a data structure having at least one data table therein;

a database controller for allowing data to be provided to and extracted from said data structure; and

a system for compressing said at least one data table, including:

4

a table modeller that discovers at least one model of data mining models with

guaranteed error bounds of at least one attribute in said data table in terms of other attributes

in different columns of said data table, ~~and~~

a model selector, associated with said table modeller, that selects a subset of said at

least one model to form a basis upon which to compress said data table to form a compressed

data table, and

a row aggregator that employs said selected subset from said model selector to

improve a compression ratio of said compressed data table via row-wise clustering.

18.   (original) The system as recited in Claim 17 wherein said table modeller employs

classification and regression tree data mining models to model said at least one attribute.

19.   (original) The system as recited in Claim 17 wherein said model selector employs a

Bayesian network built on said at least one attribute to select relevant models for table compression.

20.   (original) The system as recited in Claim 18 wherein construction of said models uses

integrated building and pruning to exploit specified error bounds and decrease model construction

time.

21.   (original) The system as recited in Claim 17 wherein said table modeller employs a

selected one of a constraint-based and a scoring-based method to generate said at least one model.

22.   (original) The system as recited in Claim 17 wherein said model selector selects said

subset based upon a compression ratio and an error bound specific for each attribute of said data

table.

Claim 23 (canceled)

24.   (original) The system as recited in Claim 17 wherein said model selector selects said

subset using a model built on attributes of said data table by a selected one of:

repeated calls to a maximum independent set solution algorithm, and

a greedy search algorithm.